

IMPLICIT STOCHASTIC OPTIMIZATION FOR DERIVING RESERVOIR OPERATING RULES IN SEMIARID BRAZIL

Alcigeimes B. Celeste

Institut für Wasserwirtschaft, Hydrologie und
landwirtschaftlichen Wasserbau
Leibniz Universität Hannover
Hannover, Germany
geimes@yahoo.com

Wilson F. Curi

Unidade Acadêmica de Física
Universidade Federal de Campina Grande (UFCG)
Campina Grande – PB, Brazil
wfcuri@df.ufcg.edu.br

Rosires C. Curi

Área de Engenharia de Recursos Hídricos
Universidade Federal de Campina Grande (UFCG)
Campina Grande – PB, Brazil
rosires@dec.ufcg.edu.br

* *Corresponding author* / autor para quem as correspondências devem ser encaminhadas

Recebido em 08/2008; aceito em 02/2009 após 1 revisão

Received August 2008; accepted February 2009 after one revision

Abstract

This paper deals with the application of Implicit Stochastic Optimization (ISO) to determine monthly operating rules for a reservoir system located in the semiarid Northeast of Brazil. ISO employs a deterministic optimization model to find optimal reservoir allocations under several possible inflow scenarios and later constructs the rules by analyzing the ensemble of these optimal releases. The operating policies provide the monthly reservoir release conditioned on the storage at the beginning of the month and the inflow predicted for the month. In addition to the classical regression analysis, this study establishes the rules by a two-dimensional interpolation strategy. After the rules are identified, they are applied to operate the system under new inflow realizations and show ability to produce policies similar to those obtained by deterministic optimization taking the same inflows as perfect forecasts.

Keywords: reservoir operation; implicit stochastic optimization; semiarid regions.

Resumo

Este artigo investiga a aplicação de Otimização Estocástica Implícita (OEI) para determinar regras de operações mensais em um sistema de reservatórios localizado no nordeste semi-árido brasileiro. OEI emprega um modelo de otimização determinística para encontrar alocações ótimas do reservatório sob vários cenários possíveis de afluências e posteriormente constrói as regras a partir da análise deste conjunto de liberações ótimas. As políticas operacionais fornecem a alocação mensal do reservatório condicionada ao armazenamento no início do mês e a afluência prevista para o mês. Além da clássica análise de regressão, este estudo estabelece as regras por meio de uma estratégia de interpolação bidimensional. Após a sua identificação, as regras são aplicadas para operar o sistema sob novas realizações de afluências e mostram habilidade para produzir políticas semelhantes às obtidas a partir de otimização determinística tendo estas mesmas afluências como previsão perfeita.

Palavras-chave: operação de reservatórios; otimização estocástica implícita; semi-árido.

1. Introduction

The extreme climatic conditions of semiarid regions such as high evaporation rates and irregular space-time distribution of precipitation require efficient management of their water resources. This is the case in most of Brazil's Northeast Region, which comprises the so-called *Polygon of Droughts*, a polygon-shaped area of about one million square kilometers characterized by periodic occurrence of droughts.

The state of Paraíba is one of the poorest of Brazil and has 98% of its area within the Polygon of Droughts. Over the last years, conflicts due to the multiple uses of water from the region's main reservoirs have increased in view of the large amount of small dams constructed upstream. Since these dams retain part of the water flowing to the reservoirs that supply the main urban areas and irrigation perimeters, the inflow to such reservoirs has reduced drastically. This has caused an increase in the frequency and duration of water rationing at important economic centers.

Apart from the water scarcity problem, the semiarid zone of Paraíba suffers with frequent disputes caused by competing water users. The region with most conflicts is located within the Piancó River Basin, in the state's *sertão* (semiarid inland area). This region contains Paraíba's largest storage reservoir, the Coremas-Mãe d'Água system, formed by two integrated reservoirs with maximum capacity of 1358 billion cubic meters.

The many and different conflicting users upstream and downstream of this system call for studies to promote its optimal operation. However, most of the studies carried out so far only evaluate the performance of the system under different inflow scenarios such as dry, normal and wet scenarios. These scenarios are used as input to simulation or deterministic optimization models and conclusions are taken from the results obtained. Although such procedure is important to evaluate the system under various conditions, the operator in practice will never be able to perfectly forecast the actual upcoming scenario. Hence, it is difficult to choose which operating policy to use (dry-, normal- or wet-scenario policy). Instead of evaluating the system's performance under pre-defined inflow scenarios it may be advantageous to find operating policies taking the uncertainties into account.

This paper deals with the application of methodologies to derive operating rules for the above-mentioned Coremas-Mãe d'Água system taking into account the uncertainties of hydrologic variables, particularly reservoir inflows, by means of a stochastic optimization method. Two types of stochastic optimization techniques have been normally used for reservoir operation: *Explicit Stochastic Optimization* (ESO) and *Implicit Stochastic Optimization* (ISO). ESO incorporates probabilistic inflow models directly into the optimization problem whereas ISO applies the Monte Carlo approach in order to take the randomness of the inflows into account. This paper tackles the problem by applying Implicit Stochastic Optimization procedures in order to obtain monthly operational policies that minimize the deficits between releases and demands. In brief, ISO uses deterministic optimization to operate the reservoir under several equally likely inflow ensembles and then examines the resulting set of optimal operating data to develop rule curves. Such curves can serve as guides to the reservoir operators to decide upon the volume of water that should be allocated at each month. After their definition, the rule curves are used to simulate the operation of the system under various possible inflow scenarios and their performance is compared to that of operations carried out by standard simulation techniques and perfect-forecast optimization.

2. Material and Methods

2.1 The Coremas Mãe d'Água system

The reservoirs Coremas and Mãe d'Água dam up, respectively, the rivers Piancó and Aguiar, located in southwest Paraíba (Figure 1). The reservoirs are connected to each other by an open canal with a capacity of 12 m³/s, and form a single lake with a water surface area of 115.6 km² at the top elevation.

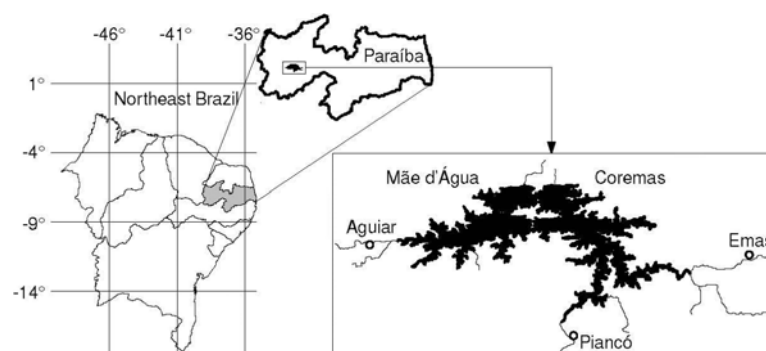


Figure 1 – Location of the Coremas-Mãe d'Água system.

The Coremas reservoir, with a 720-million-cubic-meter capacity, is responsible for hydroelectric generation and regulation of the Piancó River by means of the plant discharge that supplies all downstream demands except those of the so-called “Canal da Redenção” (Redemption Canal), which are supplied directly by the Mãe d'Água reservoir. The Coremas reservoir has two important tributaries: the Piancó River (main tributary), which controls 59.4% of the catchment area and is responsible for approximately 50% of all inflow to the system; and the Emas River, which controls 40.6% of the watershed (Lima, 2004).

The Mãe d'Água reservoir, with a 638.7-million-cubic-meter capacity, supplies an irrigation perimeter (5000 ha) at the city of Sousa (in the interior of the state) by means of the Canal da Redenção, with a total extension of 37 km. The Aguiar River is the principal tributary of the Mãe d'Água reservoir and responsible for 19.8% of all inflow to the system (Lima, 2004).

The major difficulty for the mathematical modeling of the Coremas-Mãe d'Água system is the implementation of the water transfer between both reservoirs. The difference in water surface elevation at the two lakes and the elevation of the interconnecting canal defines which reservoir is being water-fed by the other and whether there is such feeding or not, the latter case occurring when both maximum levels are below the elevation of the canal.

In this paper, the optimization methodologies are applied to define operating rules for the system as a whole (not for each single reservoir) and, therefore, the system is modeled as an equivalent reservoir. That means the active storage of the equivalent reservoir is equal to the sum of the active storages of the individual reservoirs (Table 1) as well as the inflow to the equivalent reservoir is the sum of the inflows to both reservoirs. Consequently, the obtained operating policies establish the amount of water to be allocated from the equivalent active storage (i.e., from both reservoirs) in order to meet, to the greatest extent possible, the *total* demand of the system.

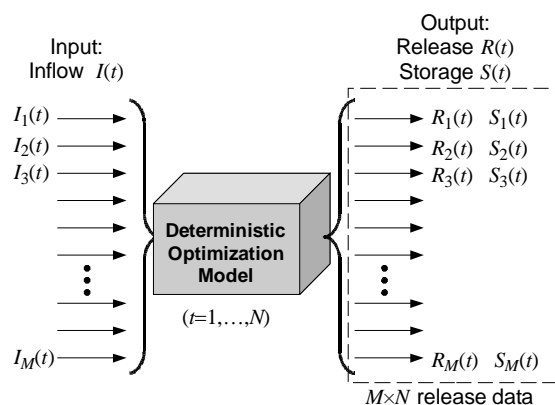
Table 1 – Characteristics of the equivalent reservoir.

	Dead Storage (hm ³)	Maximum Storage (hm ³)	Active Storage (hm ³)
Coremas	32.0	720.0	688.0
Mãe d'Água	14.8	638.7	623.9
Equivalent Reservoir	46.8	1358.7	1311.9

2.2 Implicit Stochastic Optimization

In practice, reservoir operation is usually carried out by the use of rule curves, which serve as guidance to the operators by stipulating the actions that should be taken conditioned on the current state of the system. Rule curves are typically constructed from simulation models but can also be developed by means of optimization strategies. Implicit stochastic optimization, also referred to as Monte Carlo optimization, provides a way to derive rule curves by incorporating the uncertainties of reservoir inflows.

ISO uses a deterministic optimization model to find optimal reservoir releases under several different inflow scenarios. For each inflow realization, a different operating policy is found. The set of all operating policies is then examined in order to construct reservoir release rules (Figure 2). Classically, multiple regression analysis is applied to the optimization results in order to develop operating rules conditioned on observable information such as current storage levels, previous period and/or forecasted inflows (Labadie, 2004).

**Figure 2** – Implicit stochastic optimization.

The utilization of Monte Carlo analysis for finding reservoir operating policies was first exploited by Young (1967) in a study that utilized dynamic programming applied to annual operations. The optimal releases found by the dynamic programming model were regressed on the current reservoir storage and the projected inflow for the year. The regression equation could be thus used to obtain the reservoir release at any time given the present storage and inflow conditions. Karamouz & Houck (1982) extended Young's procedure by adding one extra constraint to the optimization model specifying that the release must be within a given percentage of the release defined by the previously found operating policy.

The releases were again established by least squares multiple regression on the current-period inflow and initial storage. Kim & Heo (2000) used ISO combined with two types of linear equations for the regression analysis to define monthly operating rules for a multipurpose reservoir. Willis *et al.* (1984) devised a different approach that utilized the probability mass function of the optimal releases, conditioned on reservoir storage and inflow. Other alternatives to regression analysis are the application of artificial neural networks (Farias *et al.*, 2006; Chandramouli & Deka, 2005) and fuzzy rule-based modeling (Shrestha *et al.*, 1996) to inferring the operating rules. Unlike most previous applications of ISO, this study explores its use in a system under semiarid conditions.

In order to develop monthly rule curves, the ISO procedure follows the three basic steps described below:

1. Generate M synthetic N -month sequences of inflows;
2. For each inflow realization, find the optimal releases for all N months by means of a deterministic optimization model;
3. Use the ensemble of optimal releases ($M \times N$ data) to develop monthly operating rules.

In this study, the releases obtained by the optimization model are conditioned on the reservoir storage at the beginning of the current month and the inflow predicted for this month. One relationship (rule) is determined for each month of the year. Thus, with information of initial reservoir storage and forecasted inflow for a given month, the amount of water that should be released can be defined by the corresponding rule.

The next section outlines the generation of synthetic inflow data for the system (step 1 of the ISO procedure) and section 2.4 explains the deterministic optimization model used in step 2. For step 3, ISO models may consider various predefined forms of policies, such as linear and nonlinear polynomials, artificial neural networks, fuzzy rules, etc. (Momtahan & Dariane, 2007). The present study applies multiple nonlinear regression together with a strategy based on two-dimensional interpolation, which are described in sections 2.5 and 2.6, respectively.

2.3 Generation of synthetic inflows

The recorded inflow data at the gaging stations of Piancó and Emas were obtained from the HidroWeb data base – Hydrologic Information System of the Brazilian National Water Agency. The Piancó gaging site contains a 39-year monthly series (January 1963 to December 2001) while Emas has two series: one from November 1963 to December 1972 and another from May 1984 to December 2001.

The data of the Aguiar gaging station were taken from the work of Câmara (2000). These data belong to a 35-year series (January 1950 to December 1984) generated by the MODHAC rainfall-runoff model (Lanna & Schwarzbach, 1989). Besides, the period of unrecorded data for the Emas series (January to October 1963 and January 1973 to April 1984) was filled with MODHAC-generated data for this station also available in the work of Câmara (2000).

The historical series of inflows to the equivalent reservoir was then created by adding the series of the stations Aguiar, Piancó and Emas for the period during which all stations had data, i.e., January 1963 to December 1984. This provided a 264-month series. The box plot of these inflow data is shown in Figure 3.

Synthetic inflow series were generated by means of the classic Thomas-Fiering stochastic streamflow model (Thomas & Fiering, 1962). The Thomas-Fiering model may be viewed as a non-stationary first order autoregressive (AR1) model. It was developed to implicitly allow for the non-stationarity of seasonal data (Reddy, 1987). However, as detailed by Celeste *et al.* (2007), this model is not well suited for semiarid regions because of the problem of zeros in the data. For a better fitting, one of the model's parameters for the dry months of August and October was heuristically calibrated. Figure 4 shows the comparison of monthly mean and standard deviation between the historical inflow series and a 1000-year synthetic series generated by the model.

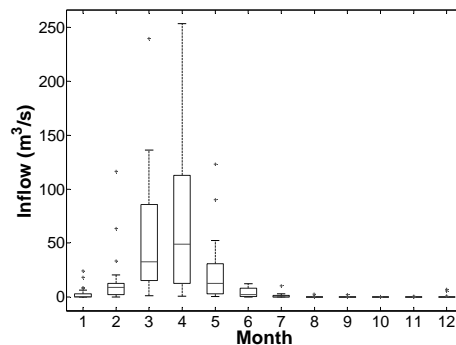


Figure 3 – Box plot of the historical inflow data.

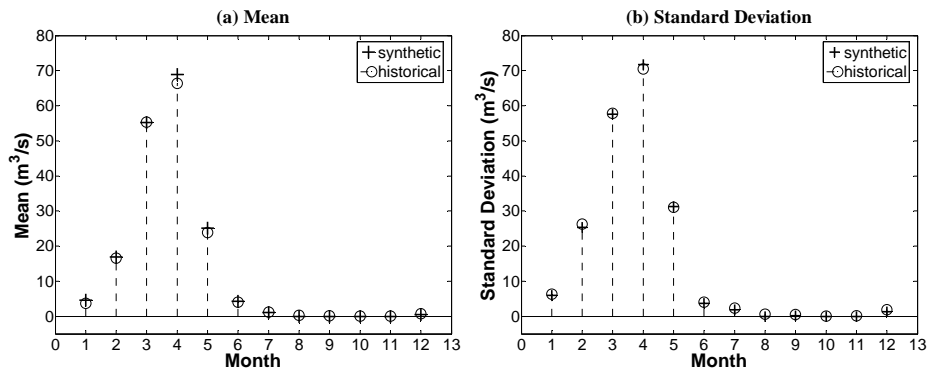


Figure 4 – Comparison of (a) mean and (b) standard deviation of historical and synthetic inflow series.

2.4 Deterministic optimization model

The deterministic operation optimization model assumes that the main objective of the operation is to find the allocations of water that best satisfy their respective demands without compromising the system. The objective function may be thus written as follows:

$$\text{minimize} \quad \sum_{t=1}^N \left[\frac{R(t) - D(t)}{D(t)} \right]^2 \quad (1)$$

where t is the month index; N is the operating horizon in months; and $R(t)$ and $D(t)$ are, respectively, release and demand in month t .

Release and storage at each period are related to inflow and spill through the continuity equation for the system:

$$S(t) = S(t-1) + I(t) - E(t) - R(t) - Sp(t); \quad \forall t \quad (2)$$

in which $S(t)$ is the reservoir storage at the end of month t (when $t = 1$, $S(t-1)$ is equal to S_0 , the initial reservoir storage); $I(t)$ and $E(t)$ are, respectively, inflow and evaporation volumes during month t ; and $Sp(t)$ is the water volume that might eventually spill from the reservoir during month t .

The physical limitations of the system define lower and upper bounds of release, storage and spill values:

$$0 \leq R(t) \leq D(t); \quad \forall t \quad (3)$$

$$S_{\text{dead}} \leq S(t) \leq S_{\text{max}}; \quad \forall t \quad (4)$$

$$Sp(t) \geq 0; \quad \forall t \quad (5)$$

where S_{dead} is the dead storage and S_{max} is the storage capacity of the reservoir.

The above optimization model represents a quadratic programming (QP) problem with linear constraints. This model was implemented in MATLAB and solved by a quadratic programming procedure that is part of the *MATLAB Optimization Toolbox*.

The monthly demand $D(t)$ of the objective function (1) was assumed to be the reservoir yield at 95% reliability. This yield was determined based on the so-called Standard Operating Policy (SOP) that specifies the total reservoir release as a function of the available water: when the water available is less than the demand, all storage water is released; and when the available water exceeds the demands, the excess is stored in the reservoir until its maximum capacity is reached and spillage starts to occur. For the reservoir equivalent to the Coremas-Mãe d'Água system, the SOP was used to construct a curve relating yield vs. reliability from the 1000-year synthetic inflow series generated by the Thomas-Fiering model. The yield at 95% reliability was found to be 12.48 m³/s.

2.5 Regression-based ISO model

The regression-based ISO model initially groups the optimal operating data found during step 2 of the ISO procedure month by month (January to December). This is done by conditioning release on reservoir storage at the beginning of the month and inflow for the month. Then, the model applies regression analysis to fit the following nonlinear equation to the data:

$$R(t) = D(t) \left[\frac{\sqrt{S(t-1)^2 + I(t)^2} - \sqrt{S_{\text{dead}}^2 + I_{\text{min}}^2}}{\sqrt{S_{\text{max}}^2 + I_{\text{max}}^2} - \sqrt{S_{\text{dead}}^2 + I_{\text{min}}^2}} \right]^m \quad (6)$$

in which I_{min} and I_{max} are, respectively, the minimum and maximum inflow values. It may be noted that the above equation is defined such that minimum storage ($S(t-1) = S_{\text{dead}}$) combined with minimum inflow ($I(t) = I_{\text{min}}$) provide no release ($R(t) = 0$); and maximum

storage ($S(t-1) = S_{\max}$) combined with maximum inflow ($I(t) = I_{\max}$) provide maximum release ($R(t) = D(t)$).

Dry months of semiarid regions usually contain many inflow values equal to zero. A month was considered dry when at least 70% of its inflow values equaled zero. For such months, the releases $R(t)$ were conditioned only on the storage $S(t)$. For dry months, the regression equation becomes:

$$R(t) = D(t) \left[\frac{S(t-1) - S_{\text{dead}}}{S_{\max} - S_{\text{dead}}} \right]^m \quad (7)$$

The task is thus to find, for each month, the value of the parameter m that fits the above equations to the data from the optimization model.

2.6 Interpolation-based ISO model

The interpolation-based ISO model uses the same data grouping as in the previous model. However, instead of fitting a curve (actually, a surface) to the data, this model uses two-dimensional interpolation in order to obtain the release given initial storage and inflow.

The interpolation is computed by means of a MATLAB routine called *gridfit* that models two-dimensional surfaces from scattered data (D'Errico, 2005). This routine can handle data with replicates and collinear points and builds a surface over a complete lattice, extrapolating smoothly into the corners. Such interpolation process might be able to better account for the high nonlinearities present in the data than the regression-based model, which uses a smooth surface to fit the data.

3. Results and discussion

The Monte Carlo process was executed over an operating horizon of 228 months (19 years). Fifty inflow sequences were generated by the Thomas-Fiering model. Each 228-month sequence was used as input to the deterministic model (1)-(5). The initial storage was set to S_{\max} . The optimization results for the first and last two years were discarded in order to avoid the influence of the boundary conditions (initial and final storages). This provided 9000 optimal release and storage values.

Release, initial storage and inflow values for January through December were grouped month by month and used to define the rules by the regression and interpolation procedures described in sections 2.5 and 2.6, respectively.

The months of August through December were found to be dry since more than 70% of their inflow values were zero. As mentioned before, the operating rules relate only release and storage for these months.

Figures 5 and 6 show the rule curves generated for April by the regression and interpolation models, respectively. Figure 7 shows the rule curve generated for the dry month of October by the regression approach.

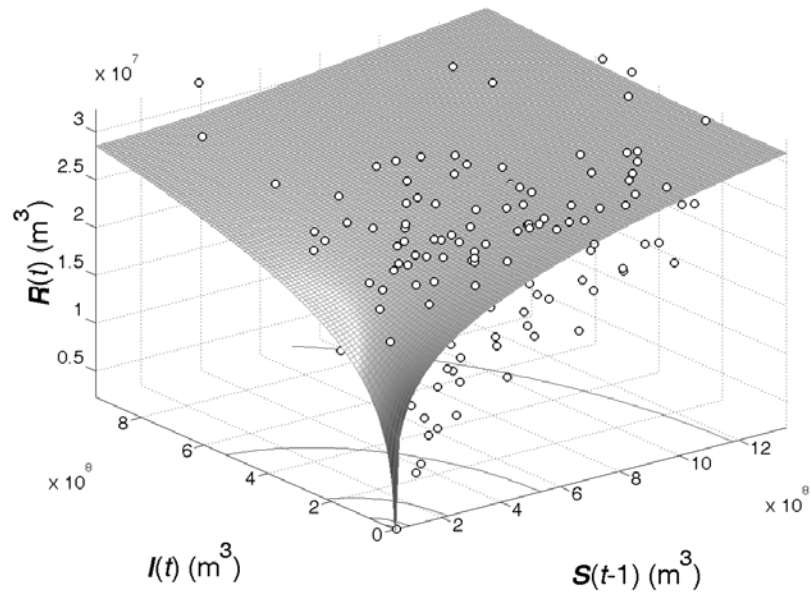


Figure 5 – Rule curve for April by the regression-based ISO approach. The circles represent the data.

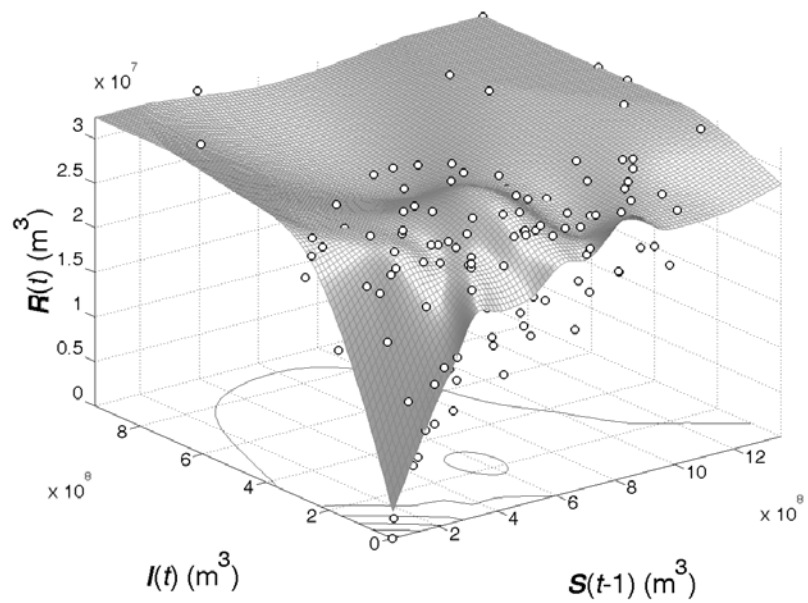


Figure 6 – Rule curve for April by the interpolation-based ISO approach. The circles represent the data.

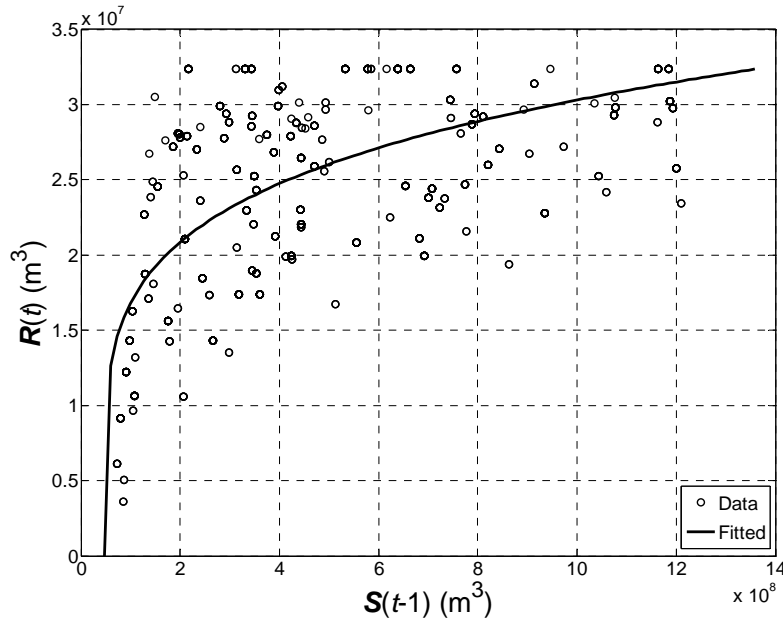


Figure 7 – Rule curve for October by the regression-based ISO approach.

After the construction of the release rules, they were applied to operate the reservoir under 10 new 19-year monthly inflow ensembles (named Sim#1, Sim#2 up to Sim#10) and the results were compared with those obtained by using the deterministic optimization model with these same inflow ensembles as perfect forecasts. The system operation under perfect forecast provides the ideal releases that should be employed for all months within the 19 years given that, in this situation, the model has knowledge of all future monthly inflows. Thus, the solutions from the perfect-forecast optimization model may be used as a benchmark. SOP-based operations were also employed for comparison.

A measure of vulnerability similar to objective function (1) was used to compare the different approaches. The values of vulnerability for all simulations are shown in Table 2. It was noted from the results that the optimization model (under perfect forecast) rationed the supply prior to shortage periods in order to mitigate the high deficits that would occur had the releases not been decreased, which happened when the SOP was used. Both ISO-based operations were capable to allocate water in a way similar to that of the optimization (i.e., the ideal operation). In average, the SOP vulnerability was 78% greater than that obtained by optimization while the deficit of the ISO rules was only 28% greater, in average, than that of the ideal operation. This indicates that the results found by the ISO-generated operating policies were very satisfactory given the fact that they had information only about the initial reservoir storage and the current-month inflow whereas the optimization model had information for the whole operating horizon and, therefore, means to define superior policies.

Table 2 – Values of vulnerability for all simulations.

	Sim#1	Sim#2	Sim#3	Sim#4	Sim#5	Sim#6	Sim#7	Sim#8	Sim#9	Sim#10
Optimization	15.26	8.26	28.17	18.03	9.38	28.84	13.82	35.07	18.59	34.59
SOP	27.35	14.89	52.44	28.12	22.96	46.46	26.29	50.52	33.95	53.44
Regression-based ISO	19.42	9.43	35.98	23.95	13.58	35.11	17.91	40.42	22.63	42.62
Interpolation-based ISO	20.90	10.40	34.89	23.96	13.66	34.65	19.49	39.34	24.53	43.90

4. Conclusions

In this study, a stochastic optimization strategy based on Monte Carlo analysis was implemented to implicitly incorporate the inflow uncertainties involved in the operation of a reservoir located in the semiarid region of Brazil. Several operating rules were derived for the system and they showed to be able to produce policies comparable to the ones obtained by deterministic optimization using perfectly forecasted inflows. In contrast to the optimization approach, however, the derived rules need to estimate only the inflow for the current month. In this way, implicit stochastic optimization shows to be advantageous against pure simulation procedures and scenario optimization and may assist operators in the management of systems in which the uncertainties cannot be ignored.

Acknowledgements

The first author wishes to acknowledge Brazil's National Council for Scientific and Technological Development (CNPq) and the State of Paraíba Research Foundation (FAPESQ) for the financial support to this research.

References

- (1) Câmara, E.P. (2000). Otimização da água do reservatório Coremas/Mãe d'Água para múltiplos usos. M.Sc. Thesis, Universidade Federal da Paraíba, Campina Grande – PB, Brazil.
- (2) Celeste, A.B.; Curi, W.F. & Curi, R.C. (2007). Análise de métodos para geração sintética de vazões em rios intermitentes. **In: Proceedings of the XXVII Simpósio Brasileiro de Recursos Hídricos**, São Paulo, CD-ROM.
- (3) Chandramouli, V. & Deka, P. (2005). Neural network based decision support model for optimal reservoir operation. *Water Resources Management*, **19**(4), 447-464.
- (4) D'Errico, J. (2005). *Surface fitting using gridfit*. Available at <<http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=8998>>.
- (5) Farias, C.A.S.; Celeste, A.B.; Kadota, A. & Suzuki, K. (2006). ANN- against interpolation-based implicit stochastic optimization for deriving reservoir operating rules. **In: Proceedings of the 15th Congress of APD-IAHR**, Madras, Chennai, India, 633-639.

- (6) Karamouz, M. & Houck, M.H. (1982). Annual and monthly reservoir operating rules generated by deterministic optimization. *Water Resources Research*, **18**(5), 1337-1344.
- (7) Kim, T. & Heo, J-H. (2000). Application of implicit stochastic optimization in the Han River basin. **In: *Proceedings of the 4th International Conference on Hydro-Science & -Engineering***, Seoul, CD-ROM.
- (8) Labadie, J.W. (2004). Optimal operation of multireservoir systems: State-of-the-art review. *Journal of Water Resources Planning and Management*, **130**(2), 93-111.
- (9) Lanna, A.E.L. & Schwarzbach, M. (1989). MODHAC – Modelo hidrológico auto-calibrável. Instituto de Pesquisas Hidráulicas, UFRGS, Porto Alegre – RS, Brazil.
- (10) Lima, C.A.G. (2004). Análise e sugestões para diretrizes de uso das disponibilidades hídricas superficiais da bacia hidrográfica do Rio Piancó, situada no Estado da Paraíba. Doctorate Thesis, Universidade Federal de Campina Grande, Campina Grande – PB, Brazil.
- (11) Momtahan, Sh. & Dariane, A.B. (2007). Direct search approaches using genetic algorithms for optimization of water reservoir operating policies. *Journal of Water Resources Planning and Management*, **133**(3), 202-209.
- (12) Reddy, P.J. (1987). *Stochastic Hydrology*. Laxmi Publications, Delhi.
- (13) Shrestha, B.; Duckstein, L. & Stakhiv, E. (1996). Fuzzy rule-based modeling of reservoir operation. *Journal of Water Resources Planning and Management*, **122**(4), 262-269.
- (14) Thomas, H.A. & Fiering, M.B. (1962). Mathematical synthesis of stream-flow sequences for the analysis of river basins by simulation. **In: *Design of Water Resource Systems*** [edited by A. Maass *et al.*], Harvard Univ. Press, Cambridge, Massachusetts, 459-493.
- (15) Willis, R.; Finney, B.A. & Chu, W-S. (1984). Monte Carlo optimization for reservoir operation. *Water Resources Research*, **20**(9), 1177-1182.
- (16) Young, G.K., Jr. (1967). Finding reservoir operating rules. *Journal of the Hydraulic Division of ASCE*, **93**(HY6), 297-321.